

# オンライン地域情報システムにおけるユーザ操作に基づく 興味度モデルの提案

小寺 里佳<sup>†</sup> 奥 健太<sup>††</sup> 北山 大輔<sup>†††</sup> 角谷 和俊<sup>†††</sup>

<sup>†</sup> 兵庫県立大学大学院環境人間学研究科 〒670-0092 兵庫県姫路市新在家本町1丁目1-12

<sup>††</sup> 立命館大学工学部 〒525-8577 滋賀県草津市野路東1丁目1-1

<sup>†††</sup> 兵庫県立大学環境人間学部 〒670-0092 兵庫県姫路市新在家本町1丁目1-12

E-mail: <sup>†</sup>nc05t069@stshse.u-hyogo.ac.jp, <sup>††</sup>oku@fc.ritsumeimei.ac.jp, <sup>†††</sup>{dkitayama,sumiya}@shse.u-hyogo.ac.jp

あらまし 近年、オンライン地図を用いた Web 上のサービスが普及している。これらのサービスでは、地図の表示領域を変更することや、カテゴリを用いて地域情報の検索を行うことが可能である。しかしながら、カテゴリ指定と地図の表示領域の変更を用いた検索操作を行うとき、カテゴリや表示領域に対するユーザの興味は変化していくと考えられる。本研究では、地域情報を検索するためのカテゴリ指定と地図操作から抽出される、ユーザの表示領域やカテゴリに対する興味を表す興味度モデルを提案する。ユーザが行ったカテゴリ指定、地図操作、地域情報の選択からカテゴリ興味度と領域興味度を算出する。また、興味度の変化から、ユーザに適する地域情報を検索するためのカテゴリや領域を推薦する手法について述べる。

キーワード 個人化, 地理情報, GIS, カテゴリ

## An User Interest Model based on User's Search Operations on the Local Information System

Rika KOTERA<sup>†</sup>, Kenta OKU<sup>††</sup>, Daisuke KITAYAMA<sup>†††</sup>, and Kazutoshi SUMIYA<sup>†††</sup>

<sup>†</sup> Graduate School of Human Science and Environment, University of Hyogo

1-1-12 Shinzaike-honcho, Himeji, Hyogo 670-0092, Japan

<sup>††</sup> Institute of Science and Engineering Ritsumeikan University

1-1-1 Nojihigashi, Kusatsu-city, Shiga 525-8577, Japan

<sup>†††</sup> School of Human Science and Environment, University of Hyogo

1-1-12 Shinzaike-honcho, Himeji, Hyogo 670-0092, Japan

E-mail: <sup>†</sup>nc05t069@stshse.u-hyogo.ac.jp, <sup>††</sup>oku@fc.ritsumeimei.ac.jp, <sup>†††</sup>{dkitayama,sumiya}@shse.u-hyogo.ac.jp

### 1. はじめに

近年、オンライン地図の表示領域変更や、カテゴリの指定操作を用いた地域情報検索サービスが普及している<sup>(注1)(注2)(注3)</sup>。これらのサービスでは、地域情報の検索結果を地図上に表示することで、地理的な位置と地域情報を関連付けて検索することができる。地域情報検索サービスの例として、図1に Yahoo! Maps の画面を示す。ユーザは検索したい地域情報が属してい

るカテゴリを指定することで、右側の地図の領域上に表示された地域情報を見て、検索を行うことができる。例えば、旅行の計画を立てる場合、ユーザはホテルや飲食店などを検索しなければならず、ホテルと交通機関との位置を考慮して地域情報を検索することが考えられる。ユーザは地図の表示領域を変更することや、カテゴリの指定を変更していくことで、適した地域情報を検索する。しかし、ユーザが検索の対象とするカテゴリや領域は変化していくことが考えられる。その場合、地域情報の検索に有効なカテゴリや地図の表示領域をユーザ自身で決定することは難しいといえる。

地域情報の表示方法としては、静的な手法と動的な手法がある。従来のサービスでは静的な手法がとられ、地図の中心から

(注1): Yahoo! Maps <http://map.yahoo.co.jp/>

(注2): Google Maps <http://maps.google.co.jp/>

(注3): Map Fan Web <http://www.mapfan.com/>



図 1 Yahoo! Maps の例

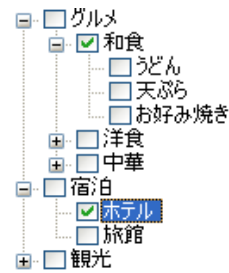


図 2 想定するカテゴリの例

の距離やカテゴリによるフィルタリングといった、ユーザの明示的な地図の領域変更やカテゴリ指定の操作によってのみ、地域情報が表示されている。しかし、従来の表示方法では、ユーザの興味を適切に反映できておらず、適した地域情報の検索が難しい場合がある。

そこで、動的に地域情報の表示を変更するために、ユーザの興味度を表現するモデルを提案する。本稿では、ユーザの操作から推定した興味度の値の変化を用いて新たなカテゴリや表示領域を推薦し、地域情報の表示を変更する手法を提示する。ユーザの興味度モデルとして、カテゴリに対する興味の強さを表すカテゴリ興味度と、領域に対する興味の強さを表す領域興味度を定義する。興味とは、そのカテゴリや領域で地域情報を検索したいというユーザの意図のことであり、ユーザが行う操作からその値を算出する。以下、2. 節において、本研究で扱う地域情報とユーザ操作について説明し、本研究の概要と関連研究について述べる。3. 節では、提案する興味度モデルについて述べる。4. 節では、興味度モデルを用いたカテゴリや領域の推薦方法について説明する。5. 節で、プロトタイプシステムの構成と推薦情報の表示について説明する。6. 節では、まとめと今後の課題について述べる。

## 2. 本研究のアプローチ

### 2.1 地域情報とユーザ操作の定義

本研究で用いる地域情報を、オブジェクト名、カテゴリ名、位置情報を持つオブジェクトと定義する。オブジェクト名とは、店舗や建物の名称のことをいう。カテゴリ名とは、そのオブジェクトが属するカテゴリのことで、オブジェクトは複数のカテゴリに属している。例えば、「和風パスタ」の店であれば、「和食」と「洋食」の両方のカテゴリに属す。位置情報とは、そのオブジェクトが存在する緯度経度のことであり、その建物や店舗の所在地を示す。本稿では、地図上に表示される地域情報のことをオブジェクトと呼ぶ。

本研究で想定するユーザの操作は、カテゴリ指定、地図操作、オブジェクト選択である。カテゴリ指定とは、ユーザがインタフェース上に表示されているカテゴリを指定することをいう。カテゴリ指定により、指定されたカテゴリに属するオブジェクトが地図上に表示される。ユーザは、オブジェクトの所在とカテゴリ情報を同時に確認しながら適するオブジェクトを検索することが可能である。オブジェクトは、指定されたカテゴリと、表示領域中の位置情報が合致した場合に地図上に表示される。例えば、ユーザが食事をする場所を探したいという場合、「和食」

のカテゴリを指定すると地図上に「和食」のオブジェクトが表示される。また、複数のカテゴリを同時に指定することも可能であるため、「和食」か「洋食」のお店に行きたいという場合にはユーザが「和食」と「洋食」両方のカテゴリを指定し、地図上にどちらかに属しているオブジェクトを表示させることが可能である。本研究で想定するカテゴリは階層構造を持ち、ユーザが上位のカテゴリを指定した場合そのカテゴリが持つ子カテゴリに属するオブジェクトが地図領域にオブジェクトとして表示される。図 2 に想定するカテゴリの例を示す。

地図操作とは、表示領域上で移動、ズームイン、ズームアウト、センタリングの操作を行うことをいう。ユーザはこれらの操作を用いて表示領域を変更し、適するオブジェクトを検索することが可能である。地図操作は、興味度の値の算出時に対象となる領域を特定するために用いる。

オブジェクト選択とは、ユーザが表示領域上に表示されたオブジェクトを選択することをいう。ユーザは、オブジェクト選択をすることで、より詳細なオブジェクトの情報を得ることが可能である。これらのユーザの操作履歴を用いて、ユーザの興味度の値を算出する。

### 2.2 研究の概要

本稿では、オンライン地図におけるユーザの地図操作とカテゴリ選択からユーザの興味の強さを表現する興味度モデルを提案する。また、興味度モデルを用いてユーザに適するカテゴリや地図の表示領域を推薦する手法を述べる。ユーザの操作として上述の地図操作、カテゴリ指定、オブジェクト選択の操作履歴を用い、興味度の値を算出する。ユーザの興味度モデルとして、カテゴリ興味度と領域興味度を定義する。カテゴリ興味度はオブジェクトが属するカテゴリに対する興味の度合いを表し、領域興味度は表示領域に対する興味の度合いを表す。カテゴリ興味度は、ユーザが指定しているカテゴリだけでなく、カテゴリ間の関係を用いて、指定していないカテゴリにも与えられる。指定していないカテゴリにも値を与えることで、ユーザが興味を持つ可能性があるカテゴリを推定することが可能である。領域興味度は、一定の区分けされた領域の単位で、その領域に対して与えられる。

ユーザに適するカテゴリや地図の表示領域を推薦する手法としては、ユーザの操作により、変化していく各興味度の値が閾値をこえたかどうかでカテゴリ、領域のどちらを推薦するかを決定する。推薦されるカテゴリや領域は、ユーザがその時点で指定をしていないカテゴリや、表示をしていない領域である。

これにより、ユーザが検索をするカテゴリや領域の候補が広がり、オブジェクト検索の補助になると考えられる。推薦の効果としては、ユーザに新たなカテゴリや領域を発見させる効果と、以前の興味を再度思い出させる効果があると考えられる。

### 2.3 関連研究

地理情報検索に関連した研究として、Paul は Web ページから抽出した地理空間データを用いて地域情報を検索する方式を提案している [10]。戸田らは、クエリと関連したキーワードと地名の近接性を考慮したランキング方式を提案している [8]。本研究はオンライン地図サービス上でのユーザの動的な操作からの興味を用いた地域情報推薦に着目した。

ユーザの嗜好を用いた情報検索に関する研究として、Teevan らはユーザの検索と文書作成の行動を基にしたユーザのプロフィールを用い、検索結果の個人化のためのアルゴリズムを提案している [9]。山本らは Web ページの検索結果に対して、ユーザが直接キーワードの強調削除の操作をすることで検索結果をリランキングするシステムを提案している [5]。Soo らは、モバイルでのユーザの行動、時間、地理的なコンテキストを含むエージェントを利用した推薦システムを提案している [6]。Martinez らは、ユーザの嗜好とプロフィールを用いて、連携システムと知識ベースシステムをかけたレストラン検索システム「REJA」を提案している [7]。これらの研究は、ユーザの嗜好に関する情報と地域情報を用いるという点で本研究と類似している。本研究では、ユーザの明示的な嗜好ではなくユーザの操作を用いて地域情報を推薦する。

以下にユーザのオンライン地図操作から抽出した意図を用いた情報検索に関連した研究として、平元らはユーザのオンライン地図に対する操作から意図を決定し、自動的にクエリを生成することで Web ページを検索するシステムを提案している [1]。手塚らは、地図上に経路を描画し、その経路に沿って Web 空間を仮想的に移動しながら地域に関連する Web ページを発見する「車窓」システムを提案している [2]。Weakliam らはユーザの地図操作を解析し、個人の操作ごとに異なる地図が面を提示するシステムを提案している [3]。これらの研究は、ユーザの意図をユーザが行った操作のパターンから抽出するという点で本研究と類似している。本研究では、ユーザの意図として興味を扱い、地図操作とカテゴリ指定から推定する。

## 3. 興味度モデル

### 3.1 興味度モデルの性質

興味度モデルとして、カテゴリに対する興味を表すカテゴリ興味度と、領域に対する興味を表す領域興味度を提案する。それぞれ、ユーザが特定のカテゴリに含まれるオブジェクトを探したいという意図、特定の領域に含まれるオブジェクトを探したいという意図の強さを表す。例えば、カテゴリ「和食」のカテゴリ興味度の値が他のカテゴリよりも高い場合、ユーザには「和食」に属するオブジェクトを探したいという意図があるといえる。また、ある駅前の領域が持つ領域興味度の値が他の領域よりも高い場合、ユーザには駅前の領域に含まれるオブジェクトを探したいという意図があるといえる。

### 3.2 カテゴリ興味度

カテゴリ興味度は、ユーザがこのカテゴリで検索をしたいという意図を表す。カテゴリ集合の定義を以下に述べる。

$$C_s = \{c_s | c_s \text{ is selected}\} \quad (1)$$

$$C_d = \{c_s | c_d \text{ is not selected}\} \quad (2)$$

$C_s$  はユーザにより指定されているカテゴリ集合、 $C_d$  は指定されていないカテゴリ集合を表し、 $c_s$  と  $c_d$  は各集合に属す任意のカテゴリを表す。

カテゴリ興味度は、値が高いカテゴリほどユーザがそのカテゴリに属するオブジェクトを探したいことを意味する。カテゴリ興味度の値は以下の式を用いて各カテゴリに与えられる。

$$CIV_{c_i,t} = CIV_{c_i,t-1} + \sum_{c_s \in C_s} Sim(c_i, c_s) \times \frac{|O_{select} \cap O_{show} \cap O_{c_s}|}{|O_{show} \cap O_{c_s}|} \quad (3)$$

式 3 において、 $CIV_{c_i,t}$  はある時刻  $t$  におけるカテゴリ  $c_i$  の興味度の値を表す。 $t-1$  は直前の操作時点を表す。また、 $c_i$  は特定のカテゴリ、 $C_s$  はユーザが選択したオブジェクトが属するすべてのカテゴリ集合を表す。 $c_s$  はユーザに指定されているカテゴリを表す。 $O_{show}$  は表示領域中表示されているオブジェクト集合、 $O_{select}$  はユーザによって選択されたオブジェクト集合を現す。また、 $O_{c_i}$  は任意のカテゴリ  $c_i$  に属するオブジェクト集合を表す。ユーザが特定のカテゴリに属するオブジェクトを多く選択した場合、そのカテゴリのカテゴリ興味度が上昇し、ユーザはそのカテゴリに属するオブジェクトを検索したいという意図を表す。 $Sim$  は、カテゴリ間の類似度のことで、0.0 から 1.0 の間の値をとり、以下の式により定義する。

$$Sim(c_i, c_j) = \frac{1}{dist(c_i, c_j)} \quad (4)$$

$dist(c_i, c_j)$  とは、階層構造をツリーと考えた場合、カテゴリ  $c_i$  から  $c_j$  へグラフ上で辿るノード数を表す。そのため、 $c_i$  と  $c_j$  が同一のカテゴリの場合、 $dist(c_i, c_j)$  は 1 を返し、 $c_j$  が親や子である場合には辿るノード数として 2 を返す。また、兄弟のカテゴリへは一度親のカテゴリを経由して辿るため、 $c_i$  と  $c_j$  が兄弟である場合には  $dist(c_i, c_j)$  は 3 を返す。 $Sim$  を用いることで、ユーザが指定しているカテゴリだけでなく、指定していないカテゴリに対しても興味度の値を伝播させることが可能である。ユーザが指定しているカテゴリ  $c_s$  と、値を算出するカテゴリ  $c_i$  が一致していれば、 $Sim$  の値は 1.0 となる。 $Sim$  は事前にカテゴリ間の類似度を算出した値を用いる。このモデルは、ユーザが興味を持つと予測されるカテゴリを特定するために使用する。 $Sim$  は他にも、Jaccard 係数を用いる場合や、店舗の説明文の文書類似度を用いる場合が考えられるが、本稿ではカテゴリツリー上の距離を用いた。

例えば、ユーザがカテゴリ「うどん」を指定して、それに属するオブジェクトを選択した。このとき、すべてのカテゴリに関して興味度の値が算出され、カテゴリ「うどん」と関係の強い「天ぷら」には興味度の値の伝播が大きい、関係の弱い「中華」には値の伝播が小さい。カテゴリ興味度の伝播により、

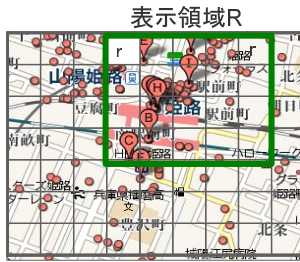


図 3 分割した領域の例

ユーザが興味を持つ可能性があるカテゴリとして「天ぷら」が候補になる．ユーザがカテゴリの指定を解除したとき，以下の式を用いて値を減少させる．

$$CIV_{c_i,t} = CIV_{c_i,t-1} - D \quad (5)$$

ここで， $D$  は定数である．カテゴリ興味度は，ユーザの表示領域変更操作により，領域中に表示されているオブジェクトが 50% 以上変化した場合に，すべてのカテゴリについて値が算出される．

### 3.3 領域興味度

領域興味度は，ユーザがこの領域で検索をしたいという意図を表す．領域興味度の値が高い領域ほど，ユーザがその領域付近でオブジェクトを探したいという意図を表す．本稿でいう領域とは，地図を一定サイズのセルに分割した  $r_i$  の集合  $R$  のことをいう．図 3 に，一定サイズに分割した領域の例を示す．表示領域  $R$  は，ユーザが表示させている地図の領域をいい，セル  $r_i$  が  $R$  中に含まれている場合と，そうでない場合で領域興味度の算出に用いる式は違う．式 6 に領域興味度の算出に用いる式を示す．

$$RIV_{r_i,t} = \begin{cases} RIV_{r_i,t-1} + \frac{|O_{select} \cap O_{show}|}{|O_{show}|} & (r_i \in R) \\ RIV_{r_i,t-1} - T & (r_i \notin R) \end{cases} \quad (6)$$

式 6 において， $RIV$  は領域興味度の値を表す． $r_i$  は分割された領域のセル 1 つを表し， $O_{show}$  は表示領域  $R$  中に表示されているオブジェクトの集合， $O_{select}$  はそのうちユーザによって選択されたオブジェクトの集合を表す．領域興味度はユーザの表示領域変更操作により，領域中に表示されているオブジェクトが 50% 以上変化した場合に，全てのセルに対して値が算出される．このとき，表示領域  $R$  中に含まれないセルは定数  $T$  で領域興味度の値を減少させる．

## 4. 推薦の判定

ユーザに対する推薦は，各興味度の値が閾値以上または以下である状態から判定される．閾値は定数で  $\alpha$ ， $\beta$ ， $\gamma$  を与える． $\alpha$  と  $\beta$  はカテゴリ興味度の値に対する閾値， $\gamma$  は領域興味度の値に対する閾値であり，推薦機能の判定に用いる興味度の値は，ユーザが指定しているカテゴリのカテゴリ興味度，ユーザが指定していないカテゴリのカテゴリ興味度，各領域のセルの領域興味度である．カテゴリ興味度をユーザが指定しているカテゴリと指定していないカテゴリで区別しているのは，ユーザが明示的に興味を持って指定をしているカテゴリと，そうでないカ

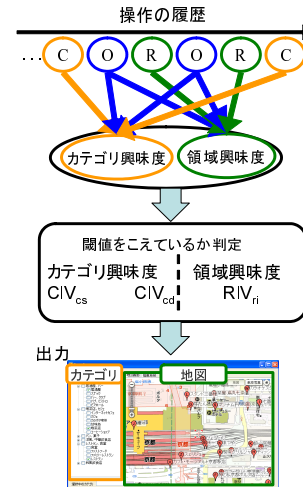


図 4 本手法の手順

テゴリを分けて扱うためである．

表 1 に興味度の状態の組み合わせをまとめる．表 1 中の I，II，III，IV は各推薦機能を表し， $\forall CIV_{c_s,t}$  はユーザが指定しているカテゴリすべて， $\exists CIV_{c_s,t}$  はそのうち任意のカテゴリが持つカテゴリ興味度， $\forall CIV_{c_d,t}$  はユーザが指定していないカテゴリすべて， $\exists CIV_{c_d,t}$  はそのうち任意のカテゴリが持つカテゴリ興味度の値を表す．また， $\forall RIV_{r_j,t}$  はすべての領域の領域興味度の値， $\exists RIV_{r_j,t}$  はそのうち特定の領域がもつ領域興味度の値を表す．

図 4 に本手法の手順を示す．図中の C，O，R はそれぞれ，ユーザが行ったカテゴリ指定，オブジェクト選択，領域変更の操作を表す．ユーザが操作を行うことで，カテゴリ興味度と領域興味度の値が算出され，興味度の値が閾値以上または以下となる状態を判定することで，カテゴリや領域をユーザに推薦情報として提示する．この手法により抽出されるユーザの興味を用いることにより，次に指定する可能性の高いカテゴリや，より多くオブジェクトを選択していた領域が推薦される．すなわちユーザはオブジェクトを検索する際，より満足するオブジェクトの閲覧数が増加することや，以前に興味を持っていた領域をすばやく確認することができると考えられる．興味度の各値が表 1 に該当する状態であるとシステムが判定したとき，カテゴリや領域が推薦情報としてユーザに提示される．

### 4.1 カテゴリの推薦

カテゴリの推薦は，表 1 中の推薦機能 I，II に該当する．これらは，ユーザがその時点で指定をしていないカテゴリを推薦する．推薦機能 I では，ユーザの領域興味度が閾値以上に上昇し，表示している領域中に存在する割合が最も高いカテゴリをユーザが指定していない場合に，そのカテゴリを推薦する．ユーザの興味度の値が，指定中のカテゴリと表示領域について閾値以上であり，かつ指定されていないカテゴリのカテゴリ興味度が閾値以下である場合，システムはこの出力を判定する．これは，ユーザがその領域に対して興味を持っているが，その領域に特徴のあるカテゴリを知らない場合，適するオブジェクトを検索するために有効であるといえる．この場合に推薦されるカテゴリ

表 1 判定に用いる興味度の状態の組み合わせ

推薦機能	指定されたカテゴリ $CIV_{c_s}$	指定されていないカテゴリ $CIV_{c_d}$	領域 $RIV_{r_j}$
I	$\exists CIV_{c_s,t} > \beta$	$\forall CIV_{c_d,t} < \beta$	$\exists RIV_{r_j,t} > \gamma$
II	$\forall CIV_{c_s,t} < \alpha$	$\exists CIV_{c_d,t} > \beta$	$\forall RIV_{r_j,t} < \gamma$
III	$\forall CIV_{c_s,t} < \alpha$	$\forall CIV_{c_d,t} < \beta$	$\exists RIV_{r_j,t} > \gamma$
IV	$\exists CIV_{c_s,t} > \alpha$	$\exists CIV_{c_d,t} > \beta$	$\exists RIV_{r_j,t} > \gamma$

りを、以下の式により定義する．

$$RecCatMax = \{c_j \mid O_{c_j} \geq \forall O_{c_i}\} \quad (7)$$

$RecCatMax$  は表示された領域中にオブジェクトが最も多く含まれているカテゴリの集合である．表示領域中に存在するオブジェクト数が最大であるカテゴリを判定し、推薦の対象とする．

例えば、ユーザが特定の地域へ旅行のために食事する場所を探す場合をあげる．旅行であれば、現地の名産品を食べたいと考えることが多い．ユーザがカテゴリ「洋食」を指定してオブジェクトを検索している時、実はその地域が「うどん」を名産にしていたとしても、ユーザは「うどん」のオブジェクトを検索することができない．そこで、システムがユーザの領域興味度の上昇を検知し、その領域内で割合の高いカテゴリ「うどん」を推薦する．ユーザに対して地域に特徴のあるカテゴリを推薦することで、検索の幅を広げ、ユーザは興味のある地域に適したオブジェクトを検索することが可能となる．

推薦機能 II では、ユーザの興味度の値が指定中のカテゴリと領域の興味度は閾値以下であり、かつ指定していないカテゴリの興味度が閾値以上に上昇した場合、閾値をこえたカテゴリを推薦する．これは、ユーザがオブジェクトを検索するとき、直接指定したカテゴリのみで適するオブジェクトを発見できない場合に有効であるといえる．この場合に推薦されるカテゴリを以下の式により定義する．

$$RecCatOver = \{c_j \mid CIV_{c_j,t} > \beta\} \quad (8)$$

$RecCatOver$  はカテゴリ興味度が閾値  $\beta$  ををこえたカテゴリを要素とする集合である．この集合の要素であり、ユーザにより指定されたカテゴリではない任意のカテゴリ  $c_j$  を判定し、推薦の対象とする．

例えば、ユーザが簡単な食事ができる場所を探す場合をあげる．ユーザはカテゴリ「喫茶店」を指定してオブジェクトを検索するが、期待する場所の付近にオブジェクトが存在しない場合には適したオブジェクトを発見することができない．そこで、システムはユーザのカテゴリに対する興味のうち、指定されていないカテゴリの興味度を算出し、他の候補となるカテゴリ「カフェ」や「コーヒーショップ」を推薦する．ユーザが指定していないカテゴリから、代替となるカテゴリを推薦することで、適した場所に存在するオブジェクトを検索することが可能である．

#### 4.2 領域の推薦

領域の推薦は、表 1 中の推薦機能 III, IV に該当する．これらは、ユーザがその時点で表示領域に含めていない領域の表示を推薦する．推薦機能 III では、ユーザがその時点で表示して

いる領域以外で、強い興味を持っている領域を推薦する．ユーザの興味度の値が、指定中のカテゴリと指定していないカテゴリでは閾値以下であり、かつ領域興味度が閾値以上である場合、システムはこの出力を判定する．これは、ユーザが新しくカテゴリを指定したとき、以前に検索していた地域に関して確認したい場合に有効であるといえる．

例えば、ユーザが大阪か京都のどちらかで食事をする場所を探す場合をあげる．ユーザが大阪でいくつかのカテゴリを用いてオブジェクトを検索し、次に京都で適するオブジェクトを検索する．京都でオブジェクトを検索する際、それまで使用していなかったカテゴリを新しく指定したとき、ユーザは以前にオブジェクトを検索していた大阪に関して同じカテゴリを用いて検索すると考えられる．しかし、地図画面を逐一大阪へ戻すことや、大阪と京都両方を含めた地図画面でオブジェクトを検索することは難しい．そこで、システムは領域興味度を考慮し、以前に操作を行ったことで領域興味度が高くなっている大阪の領域を推薦する．ユーザが以前に注目をしていない領域を確認しやすくすることで、オブジェクトの比較が容易になり、適するオブジェクトの検索が可能となる．

推薦機能 IV では、オブジェクトを検索する操作から、ユーザのカテゴリと領域に対する興味が高くなった場合に、その領域付近で興味度が高いカテゴリのオブジェクトをより多く含む領域の表示を推薦する．ユーザの興味度の値が、指定中のカテゴリ、指定していないカテゴリ、領域のすべてに閾値以上の興味がある場合、システムはこの出力を判定する．これは、ユーザが強く興味を持っているカテゴリを指定し、そのオブジェクトをより多く検索したい場合に有効であるといえる．この場合に推薦される領域を、以下の式により定義する．

$$Obj(C_l, R_n) \times \frac{|R_n \cap R_c|}{|R_n|} \geq Obj(C_l, R_c) \quad (9)$$

$Obj$  は任意のカテゴリと領域から、その領域内に存在する任意のカテゴリのオブジェクト数を返す関数である． $R_n$  は推薦の候補となる新しい領域、 $R_c$  はユーザにより直前に提示された領域である． $C_l$  はカテゴリ興味度が閾値以上であり、ユーザに指定されているカテゴリである．直前に提示された領域との重複が大きく、オブジェクト数がより多い領域を判定し、推薦の対象とする．

例えば、ユーザがカテゴリ「和食」と「洋食」どちらかで食事をする場所を探す場合をあげる．ユーザがはじめに「和食」でオブジェクトを検索した後、一度「和食」の指定を外し、「洋食」のカテゴリを指定してオブジェクトを検索した後、再度「和食」を指定した場合には、ユーザはより多くの候補となる「和食」のオブジェクトを閲覧できることが望ましい．そこで、



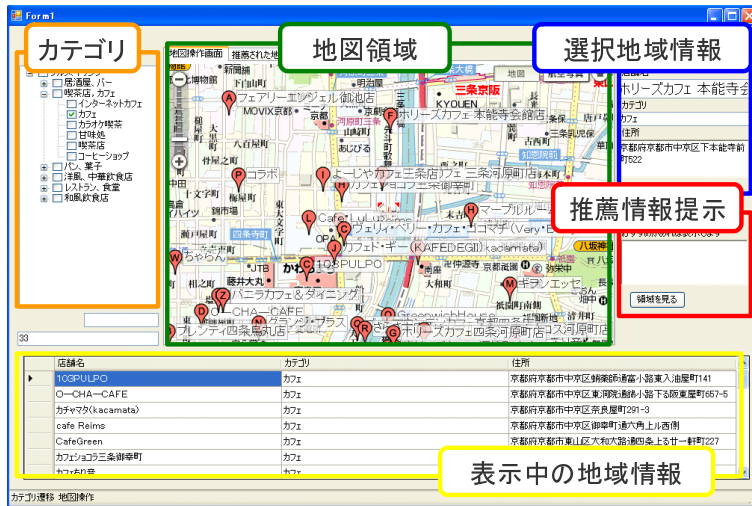


図 5 システムのインターフェース

システムはカテゴリ興味度を考慮し、興味度の高いカテゴリに属するオブジェクトをより多く含む領域を推薦する。ユーザの興味度が高いカテゴリに属するオブジェクトがより多い領域を推薦することで、検索の幅を広げ、興味のあるカテゴリで、適したオブジェクトを検索することが可能となる。

## 5. 評価実験

### 5.1 プロトタイプシステム

Visual studio 2008 を用いてプロトタイプシステムの構築を行った。オブジェクトのデータは Yahoo! 電話帳<sup>(注4)</sup>から京都市の店舗情報を取得し、3 階層のカテゴリに分類したものを利用した。図 5 にシステムのインターフェースを示す。図 5 の左側はカテゴリ部分で、ユーザがチェックボックスにチェックを入れることでカテゴリを指定することができる。カテゴリはツリー構造を持ち、ツリーを展開することで下位のカテゴリを見ることが可能である。図 5 の中央は地図の表示領域であり、ユーザは従来のオンライン地図と同様に、地図操作によって表示領域を変更することができる。地図領域ではユーザが移動、センタリング、ズームイン、ズームアウトの操作を行うことが可能である。また、オブジェクト選択は、領域上に表示されたオブジェクトをクリックすることで可能である。このとき、選択されたオブジェクトが持つ情報は、選択地域情報部分に表示される。推薦機能が判定されたとき、推薦情報提示部にテキストメッセージを表示する。図 5 の下部は表示領域中に表示されたオブジェクトが持つ情報として、オブジェクト名、カテゴリ名、位置情報としての住所を補助的に表示している。システムによって推薦する情報は、インターフェース上の推薦情報提示部にメッセージを表示することで通知する。領域が推薦された場合には、ユーザが推薦情報提示部の下にあるボタンを押すことで地図領域に推薦された領域を表示する。推薦情報提示部にメッセージを表示することで、ユーザがそのメッセージに関心を持ったときに推薦されたカテゴリを追加で指定することや、地図画面のタブ

を変更して推薦された領域を見ることができる。これは、ユーザがこれから興味を持つ可能性があるものをシステムが推薦するためである。そのため、推薦された情報に関心を持てば、領域を変更することやカテゴリ指定の追加を行うことができ、推薦情報に関心を持たなければ、そのまま検索を続行することが可能である。

ユーザが領域変更の操作を行ったことにより、表示領域が 50% 以上変化した場合に興味度を算出し、興味度の値が表 1 の状態に合致すると判定されたとき、システムはユーザにカテゴリや領域を推薦する。表示領域が 50% 以上変化したことはオブジェクト数の変化により検出する。

### 5.2 興味度モデルの評価実験

提案した興味度モデルについて、評価実験を行った。この実験から、提案するモデルがユーザが次に興味を持つ可能性があるカテゴリや、ユーザが確認に戻ると考えられる領域に対して、適切に興味度の値が与えられているかどうかを確認する、無作為に選出した大学生 10 名を被験者とし、3 つのシチュエーションを設定して提案システムの操作を行った後、アンケート調査により、次に被験者が指定する可能性があるカテゴリと、もう一度確認に戻る可能性がある領域の回答を得た。操作する領域は京都駅付近とし、被験者にはシチュエーションとして「京都で 1 人で昼食をとる場合に、食事に行く場所を探してください」、「グループで夕食をとる場合に、行く場所を探してください」、「デートで食事に行く場所を探してください」の 3 つを与えた。

被験者はシステムの操作方法の説明を受けた後これらのシチュエーションに基づいてオブジェクトを検索する操作を行い、各シチュエーションでアンケートに回答した。アンケートでは、被験者に対して「次にこのカテゴリの中で、お店を検索するために指定しようと思うカテゴリはありますか?いくつかもありません」と口頭で質問をした。また、あらかじめ京都駅付近の地図に一定間隔でメッシュを記入した地図を印刷し、その中で被験者が操作中に見ていた領域を色づけしたものを被験者

(注4): Yahoo!電話帳 <http://phonebook.yahoo.co.jp/>

表 2 興味度モデルの評価実験：カテゴリ興味度

閾値	適合率	再現率	F 値
平均値 + $\sigma$	0.31	0.23	0.26
平均値 + $1/2\sigma$	0.35	0.31	0.33
平均値	0.35	0.34	0.35
平均値 - $1/2\sigma$	0.33	0.77	0.46
平均値 - $\sigma$	0.29	0.97	0.44

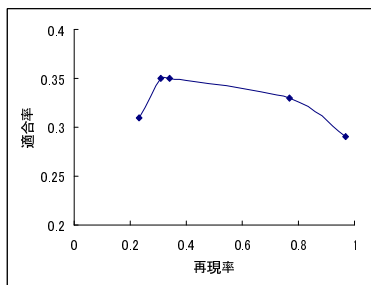


図 6 カテゴリ興味度の適合率・再現率

に提示したうえで「色をつけた部分が、さっきの検索で表示していた領域です。この中で、もう一度確認をしたい領域はありますか?」という問いを行い、確認に戻りたい領域に被験者がチェックを入れた回答を得た。被験者の回答から得られたカテゴリ数、領域数を各興味度の正解データ、システムが閾値以上で抽出したカテゴリ数、領域数をシステムの回答として適合率、再現率、F 値を算出した。推薦の対象は、最下層のカテゴリの親カテゴリ 6 つを対象とし、被験者が回答したカテゴリを子を持つ親カテゴリを正解データとして扱った。

表 2 にカテゴリ興味度の場合に、閾値を変化させたときの適合率、再現率、F 値の結果をまとめる。ここで  $\sigma$  は標準偏差を表し、平均値、標準偏差は各データごとに算出した値を用いて閾値を設定した。図 6 は表 2 の適合率と再現率をグラフに表したものである。結果から、カテゴリ興味度の閾値を平均値 -  $1/2\sigma$  とした場合に最も適切に正解データを抽出することを確認した。しかし、次に指定したいカテゴリとして、それまで指定していたカテゴリと階層が遠いカテゴリを回答する被験者も  $1/3$  程度存在した。このような被験者は、あらかじめ指定するカテゴリを想定しており、その想定に従って操作を行う途中であったといえる。そのため、システムが想定するユーザとは異なる状態であると考えられる。システムは、指定しているカテゴリとの類似を考慮して次に指定するカテゴリとして抽出するため、ユーザのカテゴリの指定し忘れた場合や、類似するカテゴリの中で別の候補を探したい場合の支援を想定している。これは、ユーザが指定しているカテゴリと類似しないカテゴリほど、ユーザが次に興味を持つとは考えられないためである。また、閲覧すべきカテゴリをすべて閲覧したとして、次に指定したいカテゴリを回答しない被験者もいた。そのようなユーザに対しても、次に指定するカテゴリの候補を提示することで、次に使用するカテゴリへの気づきを与え、オブジェクトの確認に利用できると考えられる。

表 3 に領域興味度の場合に、閾値を変化させたときの適合率、

表 3 興味度モデルの評価実験：領域興味度

閾値	適合率	再現率	F 値
平均値 + $\sigma$	0.20	0.57	0.30
平均値 + $1/2\sigma$	0.20	0.48	0.29
平均値	0.17	0.65	0.27
平均値 - $1/2\sigma$	0.05	1.00	0.09
平均値 - $\sigma$	0.04	1.00	0.09

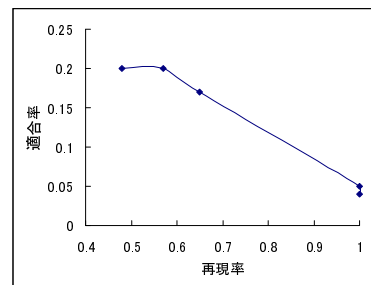


図 7 領域興味度の適合率・再現率

再現率、F 値の結果をまとめる。図 7 は各結果の適合率と再現率をグラフに表したものである。結果から、領域興味度の閾値を平均値 +  $\sigma$  とした場合に最も適切に正解データを抽出することを確認した。しかし、被験者の  $1/4$  程度は、確認したい領域がないと回答した。また、被験者が確認に戻りたい領域とは、より多くのオブジェクトが存在し、それらをもう一度確認したい領域であることや、駅など目印となる地図上のランドマーク付近であるという回答が得られた。また、領域に対する操作を観察したことから、より興味の高いオブジェクトやランドマークの付近では領域をズームインした状態でよく操作することが見られた。これらから、領域に対する興味として、よりオブジェクトを多く含む領域に対して値を多く与えることや、ズームレベルにより与える値の大きさなども考慮できると考えられる。

実験結果から、今回の実験から推薦するカテゴリは、最下層のカテゴリの抽出は難しく、また操作回数の少ないユーザに関しては興味度の値が閾値以下であり、適切に推薦すべきカテゴリや領域の抽出が難しいといえる。推薦すべきカテゴリが抽出されづらいのは、カテゴリ興味度の算出がカテゴリ間の類似度に依存していることがいえる。今回、カテゴリ間の類似度を階層関係を用いて算出したが、カテゴリに含まれるオブジェクトの個数や、パスタうどんは麺類であるため似ているというような、ユーザの感覚的な部分を考慮することや、集合の類似を求めてカテゴリの類似度を用いることが考えられる。

ツリー構造を持つカテゴリは Yahoo! カテゴリ<sup>(注5)</sup>のディレクトリ検索や、ぐるなび<sup>(注6)</sup>のレストラン検索などで一般的に用いられているが、構造により差がみられると考えられる。カテゴリの粒度が適切に作成されている場合にはカテゴリツリーを用いた手法が有用であるが、兄弟のカテゴリであるにも関わらず粒度に大きく差がある場合など、カテゴリが適切でない場合には本手法が有効に働かない可能性があると考えられる。た

(注5): Yahoo! カテゴリ <http://dir.yahoo.co.jp/>

(注6): ぐるなび <http://www.gnavi.co.jp/>

例えば、カテゴリ「和食」の兄弟として「パスタ」がある場合には、「パスタ」の粒度が小さいためユーザが「和食」と「パスタ」に対して持つ興味が当初のカテゴリの並びの時点で変化するといえ、カテゴリの粒度がユーザに影響を与えると考えられる。一方、飲食店など各カテゴリで重複するオブジェクトが多いドメインであればカテゴリ構造に依存せず、Jaccard 係数を用いた興味度の伝播を用いることが考えられる。しかし、ホテルなどカテゴリ間でオブジェクトが重複しないドメインでは、カテゴリの類似度が計算できない。このような場合には、オブジェクトが持つ価格や雰囲気といった属性をベクトルとして表現し、類似度を算出する方法も考えられる。用いるカテゴリの性質によって、カテゴリ間の関係は変化するため、どのように興味度を伝播させるべきかを検討する必要がある。

## 6. おわりに

本稿では、オブジェクト検索におけるユーザの操作に基づいた興味度モデルを提案した。興味度モデルにはカテゴリ興味度と領域興味度があり、それぞれユーザの操作を用いて値を算出する。また、興味度モデルの値の変化を用いて、ユーザの検索を補助するカテゴリや領域を推薦する手法について述べた。

評価実験により、カテゴリ興味度の場合は推薦の判定に用いる閾値を平均値 -  $1/2\sigma$  であると特定した。領域興味度の場合は、閾値を平均値 +  $\sigma$  であると特定した。これらの閾値において、ユーザが次に興味を持つ可能性があるカテゴリ、ユーザが確認に戻ると考えられる領域に対して、興味度の値が適切に与えられていることを確認した。本手法はユーザの操作から得られた興味を用いており、協調フィルタリングなどのように他者の情報を収集する必要がなく、動的に興味を算出することにより変化していくユーザの興味に対応することが可能である。

今後の課題として、興味度モデル、出力の評価、ユーザに対する推薦のインタフェース改良などがあげられる。領域興味度の算出方法として、領域に含まれるオブジェクト数の量や、ズームレベルによる値の与え方を考慮することや、カテゴリ興味度では、用いるカテゴリに依存するが、限定された地域や特定の業務に特化したカテゴリを用いることで、それらに特化した推薦を行うことも考えられる。また、カテゴリ以外でユーザが興味を持つ属性やオブジェクトの評価などを考慮することも考えられる。出力の評価実験として、プロトタイプシステムを用いて従来サービスとの性能を比較するユーザ評価の実験を行うことがあげられる。従来サービスとの比較実験では、特定の地点付近でオブジェクトを検索する場合と、2箇所の地点を比べてオブジェクトを検索する場合の2種類のシチュエーションを設定し、被験者による提案システムと従来システムの操作履歴を取得する。評価に関しては一定時間ごとで集計を行い、ユーザが選択したオブジェクト中の満足したオブジェクトと、操作回数について提案システムと従来システムでの比較を行う。また、オブジェクトのジャンルであるカテゴリ以外でユーザが興味を持つ属性や、オブジェクトの評価操作などを考慮した推薦も考えられる。

## 謝 辞

本研究の一部は、平成 22 年度科研費基盤研究 (B)(2)「ユーザの潜在的意図を用いたレス・コンシャス情報検索基盤の構築」(課題番号: 20300039) および平成 22 年度特別研究員奨励費 (21.197) によるものです。ここに記して謝意を表すものとします。

本研究の一部は、「トヨタ IT 開発センター共同研究: 地図操作とカテゴリ操作履歴マイニングによるユーザ意図抽出に関する研究」によるものです。ここに記して謝意を表すものとします。

## 文 献

- [1] 平元綾子, 角谷和俊. オンライン地図におけるユーザ操作を用いた Web 検索方式. 電子情報通信学会論文誌 (データ工学特集号) Vol.J90-D No.2, pp. 257-268, 2007.
- [2] 手塚太郎, 田中克己. クエリ・フリーな閲覧インタフェースを用いた地域情報の受動的ブラウジング. 電子情報通信学会 第 16 回データ工学ワークショップ (DEWS2005), pp. 5C-i9, 2005.
- [3] Joe Weakliam, Michela Bertolotto, and David Wilson. Implicit Interaction Profiling for Recommending Spatial Content. In Proceedings of the 13th International Symposium of Advances in Geographic Information Systems (ACM GIS 2005), pp. 285-294, 2005.
- [4] J. Goecks and J. Shavlik. Learning user's interests by unobtrusively observing their normal behavior. In Proceedings 5th International Conference on Intelligent User Interfaces (IUI'00), pp.129-132, 2000.
- [5] 山本岳洋, 中村聡史, 田中克己. ウェブ検索結果に対する語ベース適合フィードバック情報の集約によるランキング. 電子情報通信学会 第 20 回データ工学ワークショップ (DEWS2008), pp. B5-1, 2008.
- [6] Hung-Wen Tung, Von-Wun Soo. A Personalized Restaurant Recommender Agent for Mobile E-Service. In Proceedings of 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service (EEE'04), pp.1-4, 2004.
- [7] L.Martinez, R.M.Rodriguez, M.Espinilla. REJA: A Georeferenced Hybrid Recommender System for Restaurants. In Proceedings of Web Intelligence and Intelligent Agent Technologies, 2009. (WI-IAT '09). IEEE/WIC/ACM International Joint Conferences on, pp.187-190, 2009.
- [8] Hiroyuki Toda, Norihito Yasuda, Yumiko Matsuura, Ryoji Kataoka. Geographic Information Retrieval to Suit Immediate Surroundings. In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, Washington, pp.452-455, 2009.
- [9] J.Teevan, S.Dumais and E.Horvitz. Personalizing Search via Automated Analysis of Interests and Activities. In Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 449-456, 2005.
- [10] Paul Clough. Extracting Metadata for Spatially-aware Information Retrieval on the Internet. In Proceedings of the 2005 workshop on Geographic information retrieval, Bremen, Germany, pp.25-30, 2005.